Journal of Interactive Online Learning
www.ncolr.org/jiol

Volume 12, Number 1, Spring 2013
ISSN: 1541-4914

Using Learning Analytics to Predict (and Improve) Student Success: A Faculty Perspective

Beth Dietz-Uhler & Janet E. Hurn
Miami University

Abstract

Learning analytics is receiving increased attention, in part because it offers to assist educational institutions in increasing student retention, improving student success, and easing the burden of accountability. Although these large-scale issues are worthy of consideration, faculty might also be interested in how they can use learning analytics in their own courses to help their students succeed. In this paper, we define learning analytics, how it has been used in educational institutions, what learning analytics tools are available, and how faculty can make use of data in their courses to monitor and predict student performance. Finally, we discuss several issues and concerns with the use of learning analytics in higher education.

Have you ever had the sense at the start of a new course or even weeks into the semester that you could predict which students will drop the course or which students will succeed? Of course, the danger of this realization is that it may create a self-fulfilling prophecy or possibly be considered “profiling”. But it could also be that you have valuable data in your head, collected from semesters of experience, that can help you predict who will succeed and who will not based on certain variables. In short, you likely have hunches based on an accumulation of experience. The question is, what are those variables? What are those data? And how well will they help you predict student performance and retention? More importantly, how will those data help you to help your students succeed in your course? Such is the promise of learning analytics.

Learning analytics is defined as “the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” (Long & Siemens, 2011, p. 32). Learning analytics offers promise for predicting and improving student success and retention (e.g., Olmos & Corrin, 2012; Smith, Lange, & Huston, 2012) in part because it allows faculty, institutions, and students to make data-driven decisions about student success and retention. Data-driven decision making involves making use of data, such as the sort provided in Learning Management Systems (LMS), to inform educator’s judgments (Jones, 2012; Long & Siemens, 2011; Picciano, 2012). For example, to argue for increased funding to support student preparation for a course or a set of courses, it would be helpful to have data showing that students who have certain skills or abilities or prior coursework perform better in the class or set of classes than those who do not.

Learning analytics also offers the promise of more "personalized learning", which would enable students to have more effective learning experiences, among other things (Greller & Drachsler, 2012). This personalized learning experience is important in overcoming the assumption and practice of many course designers that learners start the course at the same stage and proceed through it at roughly the same pace; what Siemens refers to as the "efficient learning hypothesis" (Siemens, 2010). Without the use of performance and learning data, faculty and instructional designers are pigeon-holed into accepting this hypothesis. The use of data that is automatically collected by most LMSs allows faculty to shape how students proceed through a course. For example, Smith, Lange, and Huston (2012) found that the frequency with which students log in to their LMS, how often they engaged in the material, their pace, and assignment grades successfully predicted their performance in the course. Just as Amazon.com uses the data from our purchase history to make suggestions about future purchases, so can learning analytics allow us to suggest new learning opportunities or different courses of action to our students (Campbell, DeBlois, & Oblinger, 2007).

The purpose of this paper is to provide a brief overview of learning analytics, including various tools to track, extract, and analyze data. We will also explore its uses and applications, goals, and examples. We will discuss why individual instructors will want to make use of learning analytics. Any discussion of learning analytics is not complete without a thorough discussion of the issues and concerns with the use of this type of data.

Case Studies and Tools

There are many institutions that have made use of learning analytics to improve student success and retention. Below is a table that highlights some of these success stories. As is evident from the information in the table, many of the successful institutions have used or designed learning analytics tools that often provide a "dashboard" indicator to both students and faculty. For example, Purdue University created SIGNALS, which extracts data and provides a dashboard for both students and faculty to track student progress. Other institutions, such as UMBC, make use of a learning analytics tool built in to their institutions' LMS, which allows them to track student progress. As indicated in the third column, most of these institutions are using these data to help their students perform better in a course.

Table 1
Institutions and Learning Analytics Tools

Institution	Learning Analytic Tool	Uses of Data
University of Central Florida	EIS (Executive Information System)	Data management
Rio Salado Community College	PACE (Progress and Course Engagement)	Track student progress in course; intervention
Northern Arizona University	GPS (Grade Performance System)	Student alerts for academic issues and successes
Purdue University	Course Signals System	Student alerts for academic issues; intervention
Ball State University	Visualizing Collaborative Knowledge Work	Enhance knowledge-building work
University of Michigan	E^2Coach	Student support and intervention
University of Maryland Baltimore County (UMBC)	Blackboard LCMS	Track performance and predict student success
Graduate School of Medicine, University of Wollongong	BIRT (Business Intelligence and Reporting Tools)	Reveal continuity of care issues

There are other educational institutions that successfully use analytics to improve teaching, learning, and student success. Campbell, DeBlois, and Oblinger (2007) highlight the institutions that have achieved success by making use of various types of data to predict student success. For example, the University of Alabama used data files from first-year students to be able to develop a model of retention based on various indicators such as English course grade and total hours earned. Sinclair Community College developed their Student Success Plan (SSP) for advising and retention. Collection and analysis of these data allowed them to track students and improve student success.

In addition to these cases, there has also been success with making use of large data bases to understand student performance, with the goal of predicting success. For example, Verbert, Manouselis, Drachsler, and Duval (2012) describe various large educational datasets (e.g., dataTEL, DataShop, Mulce) that have been or can be used for learning analytic projects. Similarly, Shum and Ferguson (2012) describe several tools that can be used for social network analysis in the context of learning. Examples of such social network tools include “Mzinga”, which can be used to quantify level of participation in a network of learners, or SNAPP (Social Networks Adapting Pedagogical Practice) or “Gephi” that provide a visualization of social network participation. Dyckhoff, Zielke, Bultmann, Chatti, and Schroeder (2012) describe the development of eLAT (Learning Analytics Toolkit), which allows a user to use data to predict indicators of student success.

An Overview of Learning Analytics

As previously described, learning analytics involves the collection and analysis of data to predict and improve student success. There are vast amounts of data that faculty can make use of to help predict and improve student performance, but for our purposes, we will focus on simple forms of data that faculty have at their disposal. We are assuming that most educational institutions are using LMSs, although useful data can exist without an LMS. In this section, we describe the goals and driving forces of learning analytics, the types of data amenable to learning analytics, and how to use learning analytics.

Goals of Learning Analytics

There are a multitude of factors that have motivated interest in learning analytics. One motivating factor for the increased interest in learning analytics is the general trend for increased accountability in all levels of education. Educational institutions around the country are feeling increased pressure to account for what and how their students are learning. The pressure is even greater on online learning as these courses now have separate standards for accreditation (e.g., Ice et al., 2012). Learning analytics provides one of many methods to not only document student performance but also to provide tools that encourage the types of continuous improvement that accrediting bodies are seeking.

On a more national level, institutions of higher education are experiencing greater demands to retain students. For example, the American Graduation Initiative (Brandon, 2009) urges five million additional higher education graduates by 2020. Learning analytics can assist with this goal by providing a more personalized learning experience through the use of data to respond to students' needs (e.g., Smith, Lange, & Huston, 2012). This kind of personalization will likely lead to greater success in the classroom.

In addition to these national-level interests in learning analytics, there are more local goals that learning analytics address. These can include predicting learner performance, suggesting to learners relevant learning resources, increased reflection and awareness on the part of the learner, detection of undesirable learning behaviors, and detecting affective states (e.g., boredom, frustration) of the learner (Verbert, Manouselis, Drachsler, & Duval, 2012). As mentioned at the outset of this paper, faculty have, for the most part, relied on their intuition and hunches to know when students are struggling, or to know when to suggest relevant learning resources, or to know how to encourage students to reflect on their learning. This intuition and these hunches are not going to disappear with the advent of learning analytics, nor are the actions derived from them. Instead, learning analytics promises to make these hunches and the resulting action more data-driven and easier to detect.

Types of Data

There are vast amounts of data that educational institutions have at their disposal to help faculty meet their goals of improving students' performance and increasing retention. But, the goal of this paper is to highlight some data that faculty likely have at their fingertips that are amenable to some basic data analyses. As an example, below is a table of the types of available data in the LMS that we use at our institution (Sakai). The first column includes data that is automatically generated by Sakai. The second column is an example of the types of data that might be generated by the instructor, much of which can be stored in the LMS.

Table 2
Types of Data Available for Learning Analytics

Data Generated by LMS	**Data Generated by Instructor**
Number of Times Resource Accessed	Grades on Discussion Forum
Date and Time of Access	Grades on Assignment
Number of Discussion Posts Generated	Grades on Tests
Number of Discussion Posts Read	Final Grades
Types of Resource Accessed	Number (and Type) of Questions Asked in a Discussion Forum
	Number of Emails Sent to Instructor

The literature on learning analytics is replete with studies on the use of data such as these to predict student performance. For example, Smith, Lange, and Huston (2012) used such LMS data as login frequency, site engagement, student pace in the course, and assignment grades to predict course outcome. Macfadyen and Dawson (2012) used LMS tool use frequency (e.g., number of discussion messages read, number of discussion replies posted) to predict student achievement (course grade). Minaei-Bidgoli, Kashy, Kortemeyer, and Punch (2003) used the number of attempts at doing homework, time spent on a problem, and reading of material to predict final grades. In an interesting study, Falakmasir and Jafar (2010) examined student performance on activities that affected their final grade. They found that students' participation in a discussion forum was the best predictor of their final grades.

In a recent study, Dietz-Uhler, Hurn, and Hurn (2012) found that performance on course assignments and tests at various times in the course significantly predicted final grades in an online introductory psychology class. Specifically, they found that performance on the first two exams, a quiz on the syllabus taken before the class started, and assignments in the second half of the course accounted for 98% of the variance in final course grades. With regard to data generated from an LMS, the only variable that significantly predicted final course grade was the number of discussion board posts authored in the second half of the course. These results suggest that it is important to examine performance and behavior indicators at various points in the course in order to help students perform better in the course.

Although the amount and type of data faculty likely have available is more accessible and plentiful if using an LMS, there are other types of data that exist and can be mined beyond what is available in an LMS. For example, the widespread use of various technologies, including email, text messages, and social networks, make the amount and type of data more widespread and easier to mine (Long & Siemens, 2011).

How to Use Learning Analytics

Learning analytics can be used at various levels, including the course, curriculum, institutional, and national level. There is value in being able to leverage data analytics at all of these various levels (Dziuban, Moskal, Cavanaugh, & Watts, 2012). Although this paper is primarily focused on what faculty can do at the course level, it is wise to be mindful of these other levels.

As already mentioned, learning analytics can be used to help students succeed and to improve retention. Learning analytics can provide insights into what is happening with the learner in nearly real-time. Armed with this information, faculty can make suggestions to students that will help them succeed (Long & Siemens, 2011). For example, if a student has not read a discussion board post for a certain period of time, this may suggest to an instructor that the student needs an intervention or a nudge. Similarly, if a typically successful student suddenly performs poorly on an assignment, the instructor can intervene and seek to determine why the student performed poorly. Or, if a student repeatedly asks questions about the material or about course procedures, an instructor can examine usage data in an LMS and determine if, when, and how often the student has accessed the relevant LMS tools.

Likewise, and consistent with the goals of the national attention on assessment, learning analytics can help faculty improve teaching and learning opportunities for students (Hrabowski, Suess, & Fritz, 2011; Mattingly, Rice, & Berg, 2012). By monitoring student performance and participation in a course, as well as examining how this relates to grades, faculty can potentially spot areas of the course to improve. Such improvements in the course allow for the continual improvements that accrediting bodies are recommending. In an interesting white paper, IBM (2001) suggests ways in which educational institutions can help improve student achievement. These include:

- Monitoring individual student performance
- Disaggregating student performance by selected characteristics such as major, year of study, ethnicity, etc.
- Identifying outliers for early intervention
- Predicting potential so that all students achieve optimally
- Preventing attrition from a course or program
- Identifying and developing effective instructional techniques
- Analyzing standard assessment techniques and instruments
- Testing and evaluation of curricula

Each of these applications is amenable to learning analytics and can generally be accomplished by gathering data from an LMS or from instructor records. In addition, being able to mine the type of data necessary to achieve these goals will go far in helping to improve student success and increase retention as doing so is likely to optimize student learning experiences (Olmos & Corrin, 2012).

May (2011) suggests that learning analytics can be both descriptive and predictive. From a descriptive perspective, learning analytics can help us answer such questions as: “What happened?”, “Where was the problem?”, and “What actions are needed?”. Learning analytics can also help us to predict and prescribe by answering such questions as: “Why is this happening?”, “What if these trends continue?”, “What will happen next?”, and “What is the best that can happen?”. This approach is also consistent with the five stages of the use of learning analytics in higher education suggested by Goldstein and Katz (2005): data extraction, performance analysis, what-if decision support, predictive modeling, and automatic response triggers.

Benefits of Using Learning Analytics

We have highlighted the various reasons for wanting to use learning analytics. Among these are to improve student success, increase retention, and improve accountability. But as an individual faculty member, why might you want to make use of learning analytics on a smaller, course-level scale? In this section, we highlight several of the benefits and advantages of using learning analytics.

Long and Siemens (2011) describe a multitude of benefits of using learning analytics for higher education. Several of these benefits are focused on an administrative level, such as improving decision-making and informing resource allocation, highlighting an institution's successes and challenges, and increasing organizational productivity. From the perspective of a faculty member, they suggest that learning analytics can help faculty identify at-risk learners and provide interventions, transform pedagogical approaches, and help students gain insight into their own learning. Having data at hand and knowing what to do with it can allow us to realize these benefits. For example, if we learn (from a correlational analysis) that student performance on certain activities is not related to final grades, then we might consider modifying these activities. Similarly, we can use data from an LMS to build a model of successful student behaviors, which might include the frequency of LMS tool use, frequency of accessing discussion board posts, and the number of times taking quizzes. If we can build a model of successful student behaviors, then we can encourage (with data!) our students to engage in these behaviors. Alternatively, we can also identify at-risk students as ones who deviate from this model.

Likewise, Greller and Drachsler (2012) suggest that learning analytics can help faculty by informing them of the gaps in knowledge displayed by their students. Understanding these knowledge gaps can help faculty focus their attention on particular students or pieces of information. Of course, for institutions such as Purdue or Rio Salado that have performance dashboards in their LMSs, students can constantly monitor their progress and determine how they are performing, so there are benefits at the student level as well.

Concerns/Issues

There are a number of issues and concerns that should be highlighted in any discussion of learning analytics. At the forefront of these issues is the role of pedagogy in data analytics. It seems clear that the pedagogy should drive learning analytics and not necessarily the converse (Greller & Drachsler, 2012). Campbell et al. (2007) provide a list of the issues and concerns that must be addressed before implementing any program or course of action on learning analytics. Some of these concerns include:

- Big brother: It may be threatening to some students and faculty to know that someone can "watch" and track all that they do.
- Holistic view: There is a concern that any data set, no matter how comprehensive, cannot take in to account other issues, such as interpersonal ones.
- Faculty involvement: Faculty need to be involved in order for learning analytics to have its greatest impact.
- Obligation to act: Are faculty and institutions obligated to use data to increase the probability of student success?

In an interesting article on the potential harmful effects of learning analytics, Dringus (2012) argues that learning analytics (in online courses), must get meaningful data, have transparency, yield good algorithms, lead to effective use of the data, and inform process and practice. Without attending to these minimal requirements, Dringus argues that learning analytics can be harmful.

Other issues that are frequently highlighted in discussions of learning analytics include profiling and how learning-analytics data will be used. Specifically, there is a danger of creating a profile of successful and unsuccessful students. More importantly, there is concern that a profile creates a set of expectations for the student and faculty (Campbell et al., 2007). Of course, students and faculty already have expectations – the issue is that learning analytics might add a set of data-driven expectations.

Data privacy and the use of data are also strong concerns of the use of learning analytics. There are legal and ethical issues, such as FERPA, that need to be addressed before faculty or institutions can make use of some student data (Campbell et al., 2007; Greller & Drachsler, 2012). Similarly, there is the issue of who the data belong to. Once the data have been warehoused, can anyone have access to it (Campbell et al., 2007; Greller & Drachsler, 2012)?

Finally, there is the issue of whether or not we are really measuring student learning, or are we just attempting to boost student retention and course completion (Watters, 2012). If one considers the types of data that are mined for learning analytics, such as the number of course tools accessed in an LMS, or the number of posts "read" on the discussion forum, are these really proxies for learning? This is not to suggest that learning analytics cannot boost learning, but we need to be clear about what we are measuring and predicting.

Conclusion

It seems clear that learning analytics is gaining momentum and is likely here to stay (e.g., Horizon Report, 2012). There are many benefits to learning analytics; most notably that it can inform how we help our students succeed. In educational institutions, we have an enormous amount of data at our disposal. Our ability to harness this data and use it to inform what we do in the classroom, whether face-to-face or online, is at the heart of learning analytics. Institutions such as Purdue University, Rio Salado Community College, and University of Michigan are blazing the trail and demonstrating the vast benefits of learning analytics for students and faculty. But, as studies have indicated (e.g., Abdous, He, & Yen, 2012; Dietz-Uhler & Hurn, 2012; Falakmasir & Jafar, 2010; Minaei-Bidgoli, Kashy, Kortemeyer, & Punch, 2003), individual faculty can make use of the data they have available in their courses to affect change and improve student success. These efforts, although seemingly "small-scale", can have a large impact on student success.

References

Abdous, M., He, W., & Yen, C. (2012). Using data mining for predicting relationships between online question theme and final grade. *Educational Technology & Society, 15*(3), 77-88.

Brandon, K. (2009). Investing in Education: The American Graduation Initiative. *Office of Social Innovation and Civic Participation*. Retrieved November 22, 2012, from http://www.whitehouse.gov/blog/Investing-in-Education-The-American-Graduation-Initiative

Campbell, J. P., DeBlois, P. B., & Oblinger, D. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE Review, 42*(4), 40-57.

Dietz-Uhler, B., Hurn, J.E., & Hurn, B. (2012). *Making use of data in an LMS to predict student performance: A learning analytics investigation*. Unpublished manuscript.

Dringus, L. P. (2012). Learning analytics considered harmful. *Journal of Asynchronous Learning Networks, 16*(3), 87-100.

Dyckhoff, A. L., Zielke, D., Bultmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and implementation of a learning analytics toolkit for teachers. *Educational Technology & Society, 15*(3), 58-76.

Dziuban, C., Moskal, P., Cavanaugh, T., & Watts, A. (2012). Analytics that inform the university: Using data you already have. *Journal of Asynchronous Learning Networks, 16*(3), 21-38.

Falakmasir, M. & Jafar, H. (2010). *Using educational data mining methods to study the impact of virtual classrooms in e-learning.* Paper presented at the Proceedings of the 3rd International Conference on Educational Data Mining, Pittsburgh, PA. Retrieved from http://alum.sharif.edu/~falakmasir/images/EDMPaper.pdf

Greller, W. & Drachslrer, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society, 15*(3), 42-57.

Goldstein, P. & Katz, D. (2005). Academic analytics: The uses of management information and technology in higher education. *EDUCAUSE Center for Applied Research*. Retrieved from http://net.educause.edu/ir/library/pdf/ekf/ekf0508.pdf

Hrabowski, F. A., Suess, J. J., & Fritz, J. (2011). Assessment and analytics in institutional transformation. *EDUCAUSE Review Online*. Retrieved from http://www.educause.edu/ero/article/assessment-and-analytics-institutional-transformation

Ice, P., Diaz, S., Swan, K., Burgess, M., Sharkey, M., Sherrill, J., Huston, D., & Okimoto, H. (2012). The PAR framework proof of concept: Initial findings from a multi-institutional analysis of federated postsecondary data. *Journal of Asynchronous Learning Networks, 16*(3),63-86.

IBM. (2001). *Analytics for achievement: Understand success and boost performance in primary and secondary education*. Retrieved from http://public.dhe.ibm.com/common/ssi/ecm/en/ytw03149caen/YTW03149CAEN.PDF

Johnson, L., Adams, S., and Cummins, M. (2012). The *NMC Horizon Report: 2012 Higher Education Edition.* Austin, Texas: The New Media Consortium.

Jones, S. J. (2012). Technology review: The possibilities of learning analytics to improve learner centered decision making. *The Community College Enterprise, Spring, 2012*, 89-92.

Long, P. D. & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education.

EDUCAUSE Review Online. Retrieved from http://www.educause.edu/ero/article/penetrating-fog-analytics-learning-and-education

Macfadyen, L. P. & Dawson, S. (2012). Numbers are not enough: Why e-learning analytics failed to inform an institutional strategic plan. *Educational Technology & Society, 15*(3), 149-163.

Manaei-Bidgoli, B., Kashy, D. A., Kortmeyer, G., & Punch, W. (2003). *Predicting student performance: An application of data mining methods with an educational web-based system (LON-CAPA)*. Paper presented at the 33rd ASEE/IEEE Frontiers in Education Conference, Boulder, CO. Retrieved from http://lon-capa.org/papers/v5-FIE-paper.pdf

Mattingly, K. D., Rice, M. C., & Berge, Z. L. (2012). Learning analytics as a tool for closing the assessment loop in higher education. *Knowledge Management & E-Learning: An International Journal, 4*(3), 236-247.

May, T. A. (2011). Analytics, University 3.0, and the future of information technology. *EDUCAUSE Review Online*. Retrieved from http://net.educause.edu/ir/library/pdf/ERM1159.pdf

Olmos, M. & Corrin, L. (2012). Learning analytics: A case study of the process of design of visualizations. *Journal of Asynchronous Learning Networks, 16*(3), 39-49.

Picciano, A. G. (2012). The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Networks, 16*(3), 9-20.

Shum, S. B. & Ferguson, R. (2012). Social learning analytics. *Educational Technology & Society,15*(3), 3-26.

Siemens, G. (2010). What are learning analytics. *ELEARNSPACE: Learning, networks, knowledge, technology, community.* Retrieved from http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/

Smith, V. C., Lange, A., & Huston, D. R. (2012). Predictive modeling to forecast student outcomes and drive effective interventions in online community college courses. *Journal of Asynchronous Learning Networks, 16*(3), 51-61.

Verbert, K., Manouselis, N., Drachsler, H., & Duval, E. (2012). Dataset-driven research to support learning and knowledge analytics. *Educational Technology & Society, 15*(3), 133-148.

Watters, A. (2012). Learning analytics: Lot of education data...Now what? *Hack Education*. Retrieved from http://hackeducation.com/2012/05/04/learning-analytics-lak12/